10/509640
EU03/3585
Rec'd PCT/PTO 28 SEP 2004

PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS
The Patent Office
INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

RECD 29 JUL 2003
WIPO PCT

# PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., or PLC.

der the Companies Act does not constitute a new legal entity but merely any to certain additional company law rules.

Signed

Dated 11 June 2003



BEST AVAILABLE COPY

**Patents Form 1/77**

Pat[illegible]s Act 1977
(Rule 16)


1/77

30MAR02 E707495-5 D02718
P01/7700 0.00-0207392.2

# Request for grant of a patent

*(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)*

**The Patent Office**

Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference
SSA/AH/P8663GB

2. Patent application number
*(The Patent Office will fill in this part)*
0207392.2
28 MAR 2002

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*
Neural Technologies Limited
Ideal House
Bedford Road
Petersfield
Hampshire
GU32 3QA

Patents ADP number *(if you know it)*
7100118001

If the applicant is a corporate body, give the country/state of its incorporation
England and Wales

4. Title of the invention
A Configurable Data Profiling System

5. Name of your agent *(if you have one)*
DAVID KELTIE ASSOCIATES

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*
12 NEW FETTER LANE
LONDON
EC4A 1AG

Patents ADP number *(if you know it)*
4014502006

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

| Country | Priority application number *(if you know it)* | Date of filing *(day / month / year)* |
|---|---|---|

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

| Number of earlier application | Date of filing *(day / month / year)* |
|---|---|

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*
*a) any applicant named in part 3 is not an inventor, or*
*b) there is an inventor who is not named as an applicant, or*
*c) any named applicant is a corporate body.*
*See note (d))*
Yes

BEST AVAILABLE COPY

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

| | |
|---|---|
| Continuation sheets of this form | |
| Description | 5 |
| Claim(s) | |
| Abstract | |
| Drawing(s) | 2+2 |

10. If you are also filing any of the following, state how many against each item.

| | |
|---|---|
| Priority documents | |
| Translations of priority documents | |
| Statement of inventorship and right to grant of a patent *(Patents Form 7/77)* | |
| Request for preliminary examination and search *(Patents Form 9/77)* | |
| Request for substantive examination *(Patents Form 10/77)* | |
| Any other documents *(please specify)* | |

11. I/We request the grant of a patent on the basis of this application.

Signature David Keltie Associates Date 28/3/02

12. Name and daytime telephone number of person to contact in the United Kingdom ANNA HAMADYK 020 7583 6000

**Warning**

*After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.*

**Notes**

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

BEST AVAILABLE COPY

# A Configurable Data Profiling System

## Introduction

Fraud is a serious problem in modern telecommunications systems, and can result in revenue loss by the telecommunications service provider, reduced operational efficiency, and increased subscriber churn. In the highly competitive telecommunications sector, any provider that can reduce the revenue loss resulting from fraud – either by its prevention or early detection – has a significant advantage over its competitors.

Differences in networks and services exist not only on an international level, but also between operators in individual countries. For example, different operators may specialise in only mobile or landline services, each of which have unique fraud characteristics, and thus require different fraud detection engines. Similarly, different countries may have different standards for the B-number (destination number) partitions that distinguish different types of services, thus requiring modifications to B-number sensitive components of a fraud detection engine.

For example, telephone networks in the UK prefix the numbers of premium rate services with 0898 and freephone services with 0800. Most fraud detection systems in operation in the UK therefore consider high volumes of calls to numbers starting with 0898 to be more suspicious than to numbers starting with 0800 because the high cost of calls to premium rate services makes them an attractive target for fraudsters. If UK-based fraud detection engines are transferred to other countries, they will need to be modified to account for the fact that the prefixes that indicate premium rate and freephone services are different.

The patterns that characterise fraudulent behaviour also change with time, not least in response to a telco's attempts at detection and prevention. A fraud detection system therefore needs to be highly configurable so that it can easily be adapted to the requirements of different networks and operators, and to incorporate information about new types of fraud as they emerge. Such configuration must be possible without modification to the fraud detection software, as the development, testing, and validation processes are too expensive and time consuming to be repeated often enough to keep fraudsters in check.

Most fraud detection systems identify fraud by building profiles of the behaviour of particular entities in a network based on a pre-defined, hard coded set of features, such as average call duration, or the percentage of calls to international numbers, which are measured over fixed or variable time periods (see, for example, WO0141469). Such systems cannot be modified to detect new fraud types, or to operate in environments where the pre-defined feature set is not effective without software modifications. This document describes a system for constructing and processing behavioural profiles that are well suited to fraud detection, that is highly flexible, and can be configured without requiring changes to the underlying software engine.

BEST AVAILABLE COPY

**Description**

A high level representation of the system is given in figure 1. The system consists of three functional modules, a pre-processing module (11), a profiling module (12), and a post-processing module (13). The details of the functioning of each of these can be configured at runtime using information supplied to the system from some external source (such as a graphical user interface (GUI), configuration file, or configuration data stream). To maximise efficiency, the system is event driven, and only performs processing when changes occur in its inputs (other than those originating in the post-processing module).

The pre-processor (11) receives information from an external data stream (14) (which can, for example, contain information from event data records (EDRs), which are generated whenever a call is made), customer and business data, or information output by the post-processing module (13)). An EDR is a collection of data that describe an event that has occurred in or on a network. Events such as the start or end of telephone calls result in the creation of EDR's that include information such as the call's start time, its duration, cost, the number dialled, etc.

The data stream (14) can contain multiple substreams, the contents of which can be unrelated and can change in unrelated ways. For example, a data stream may contain a customer data substream, and an EDR substream. The contents of the customer data substream would only change when customer details change (for example, as a result of a change of address), while the contents of the EDR substream would change with every call. The pre-processor can perform a mixture of runtime-configurable linear and non-linear transformations of its inputs. These transformations can consist of mathematical and logical functions and rules, which have access to external databases. The results of these transformations are called 'profiling features' and each consists of a numeric scalar or a string. For example, a list of 'hot' destinations (numbers that are frequently called by fraudsters) can be stored in an external database (18). A profiling feature can then be created that indicates whether an EDR represents a call to one of the listed numbers by assigning to it a value of one if the B-number in the EDR matches one of the listed hot destinations, and zero otherwise.

The pre-processing module supports the creation of intermediate variables, which persist only while the pre-processing module is active, and can be used to store the intermediate results of calculations. This important feature improves the efficiency of the pre-processing module by allowing intermediate results – which may be common to several functions within the module – to be calculated once and used many times. More permanent storage is available to the system in the profiling module (12). Different functions can be applied to each of the pre-processor's inputs and combinations thereof. Linear functions can be used to allow information to pass through the pre-processing module unchanged. The pre-processing module outputs profiling features (15) that are used by the profiling module (12) to construct a profile of an entity's behaviour. Each profiling feature (15) can be flagged as changed or unchanged by the pre-processor (11) according to its configuration, and the profiling module (12) is updated only if at least one of the profiling features (15) is flagged as changed. This improves the efficiency of the invention because the pre-processor configuration can prevent the entire system being updated if changes in its input (14) are not considered significant by marking all profiling features as unchanged.

BEST AVAILABLE COPY

The profiling module (12) is shown in more detail in figure 2. It summarises the behaviour of each profiling feature (21) over a time window by dividing it into a number of non-overlapping time slots (22) of configurable length (the figure shows a profile based on two hour slots). When a set of profiling features (15) is presented to the profiler (12), they are entered into the profile according to the profiler's configuration. The profiler can be configured in a variety of ways, including storing the feature information in the slot during which the event that caused the profile to be updated started or ended, or in every slot during which the event was in progress. If the selected update mechanism goes beyond the end of the last (most recent) slot in the time window, it "wraps around" to the first (oldest) slot and overwrites the information within it. Event messages are generated by the profiling module (12) whenever significant events (such as the time window wrapping around, or when the module receives its first input) occur within it. These messages can be used within the profiler configuration to trigger specific rules, or passed to the post processing module (13).

When entering new information into time slots, the profiler (12) can be configured so that it either overwrites that already present in the selected slot(s), or is added to it. In addition to these basic features, almost any way of changing the information in the profile can be implemented using the feedback loop between the post-processor (13) and the pre-processor (11), and the rules and functions supported by those modules. For example, information in a time slot can effectively be multiplied by new information by taking the logarithm of the new information in the pre-processor (11), adding it to the contents of the selected slot(s) and forming the exponential of the slot contents in the post-processing module (13). In addition to the slot-based time window, the profiler (12) also contains an area of scratchpad memory (24) for the general storage of quantities in a way that is independent of the start and end times of the events with which they are associated. Sections of the scratchpad memory can be designated as volatile, which means that they only exist while the system is active, and are not stored in the application database with the rest of the profile. They provide temporary storage within the system, and can be used to transfer data from the pre-processing module to the post-processing module unchanged, effectively bypassing the profiler.

The post-processing module (13) is essentially the same as the pre-processing module, except that it operates on the profiled feature information (16), which consists of the contents of the profiler time slots (23) and the contents of the scratchpad memory (24). The role of the post-processing module is threefold – firstly, to process information in the profiler (12) that is to be fed back to the pre-processor (11), secondly, to process information ready for presentation to other components in the fraud detection system, and thirdly, to perform some fraud detection directly. This latter goal can be achieved by configuring rules, for example, within the post-processing module to search for suspicious characteristics within the profiled features (16). The output of the entire profiling system (17) thus consists of post-processed profiling information, and, potentially, fraud indications. Post-processed profiling information is typically subject to further processing - for example, by the application of rules, scorecards, change detection algorithms and other statistical analyses - in order to identify suspicious behaviour. Fraud indications are typically sent to another layer of processing for further analysis. Like the pre-processor (11), each output of the post-processor (13) can individually be flagged as changed or unchanged according to the post-processor's configuration. This allows the invention to be used within a larger event-driven system, and to cause updates to it to be triggered only when significant events occur.

BEST AVAILABLE COPY

A simple example of how the system could be applied in practice is shown in figure 3. In this example, the system is configured to calculate the total cost of calls made within one hour periods. Note that there are numerous alternative ways of using the present system for this purpose, all of which create a unique instance of the system for each user. Whenever a call is made, an EDR (31) is generated within the telecommunications network, and passed to the instance of the system dedicated to the user that made the call. The call cost information – which is usually contained in the EDR (31) – can be passed through the pre-processor (32) to the profiling module unchanged (33), creating a numeric profiling feature with a value equal to the cost of the current call (34). (If call cost information was not available in the EDR, the pre-processor could be configured to calculate call cost estimates using a series of rules and equations, based on information on the duration of a call, the destination number, and a lookup table of call charges stored in a database.) The pre-processor would be configured to flag the call cost profiling feature (34) as changed every time a new EDR (31) arrives to ensure that the profiling module (33) is updated for every call.

The call cost profiling feature (34) can be used in several ways. In this example, it is simply accumulated within the profiler time slots (35) corresponding to the call start times to form a measure of the total cost of all calls made by the user that were started within each slot. In more sophisticated realisations, recursive estimates of summary statistics of call cost (such as its mean and variance) can be formed in the scratchpad storage by appropriately formulated pre-processor (32) rules. The post-processor (36) is, in this example, configured to output the contents of the first profiler slot earlier than that updated by the current call, and to mark that output as changed only if the current call was the first in a new slot. For example, in figure 3, the call occurs at 8:32am, which falls in the second slot, so the contents of the first slot – in this case 3.97 – will be output by the post-processor (36). This output will only be marked as changed if the current call is the first one to fall within the second slot. Since, in figure 3, several calls have already been recorded in the second slot, the post-processor's output (37) will be marked as unchanged. More complex post-processor (36) configurations are also possible, so that it can, for example, contain rules that generate alerts if the total call cost within an individual timeslot exceeds a predefined threshold, or if the cost of an individual call exceeds some function of mean and variance cost estimates. In the latter example, individual call costs can be passed to the post-processor (36) using volatile scratchpad storage.

## Summary

This document describes a fully configurable system for performing time-based profiling of data streams. It is believed that novel aspects of the system include:

- A runtime configurable pre-processing component that supports a wide range of logical and mathematical functions,
- A runtime configurable profiler that contains general purpose scratchpad memory
- A runtime configurable post-processing component that supports the same set of functions as the pre-processing component,
- A runtime configurable feedback loop that allows elements of the profile to be used as inputs to the pre-processing module.

BEST AVAILABLE COPY

Generally, the runtime configurability of the system is of prime importance, because:

- It allows the fraud detection capabilities of the system to be modified without the underlying software engine being changed, recompiled, tested, and validated. This,
  - Reduces the time required to incorporate new fraud detection algorithms into the fraud detection engine, and hence helps to keep fraudsters in check,
  - Reduces the risk of potentially serious bugs being introduced into the fraud detection engine. Since the configuration environment can be carefully controlled, it is easier to guarantee that changes to the configuration do not create serious bugs than is possible with changes to the underlying software engine, and
  - Can allow non-programming personnel to configure the fraud detection engine, provided that the configuration mechanisms are suitably formulated.

Notes

The operational profile is a highly configurable system for extracting synoptic profiles of a subscriber's behaviour over a fixed period of time. Competing products offer similar systems, but lack the runtime configurability of the Minotaur OP. This document describes a system similar to the Minotaur OP.

BEST AVAILABLE COPY

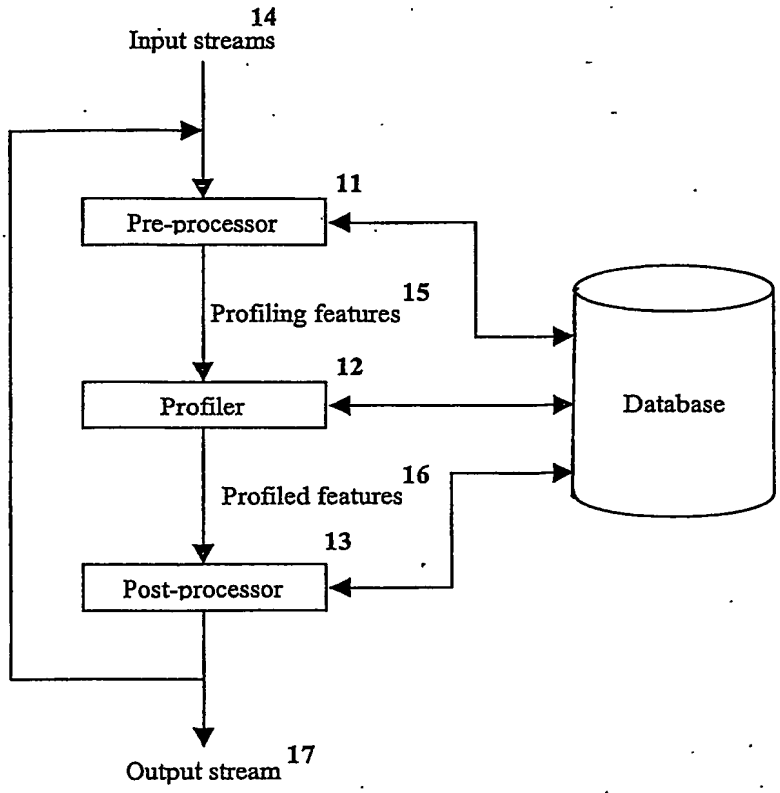


Fig. 1

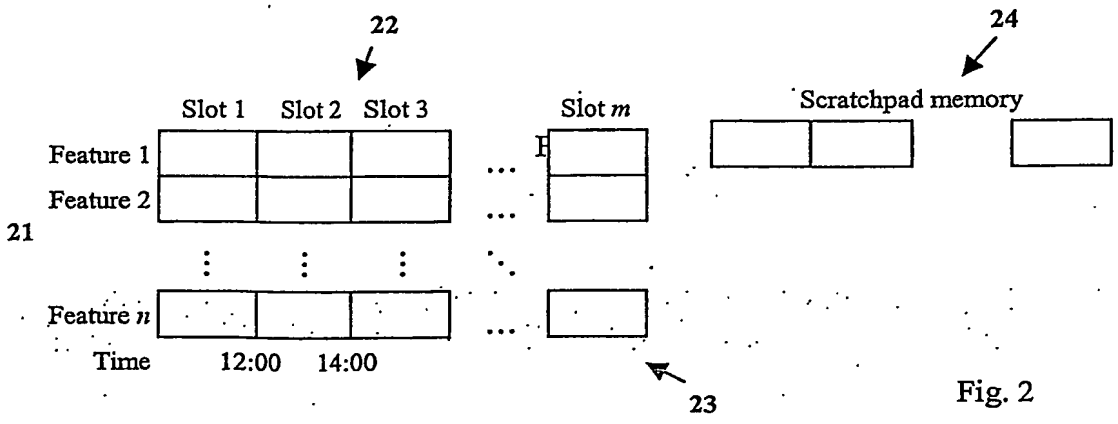


Fig. 2

BEST AVAILABLE COPY

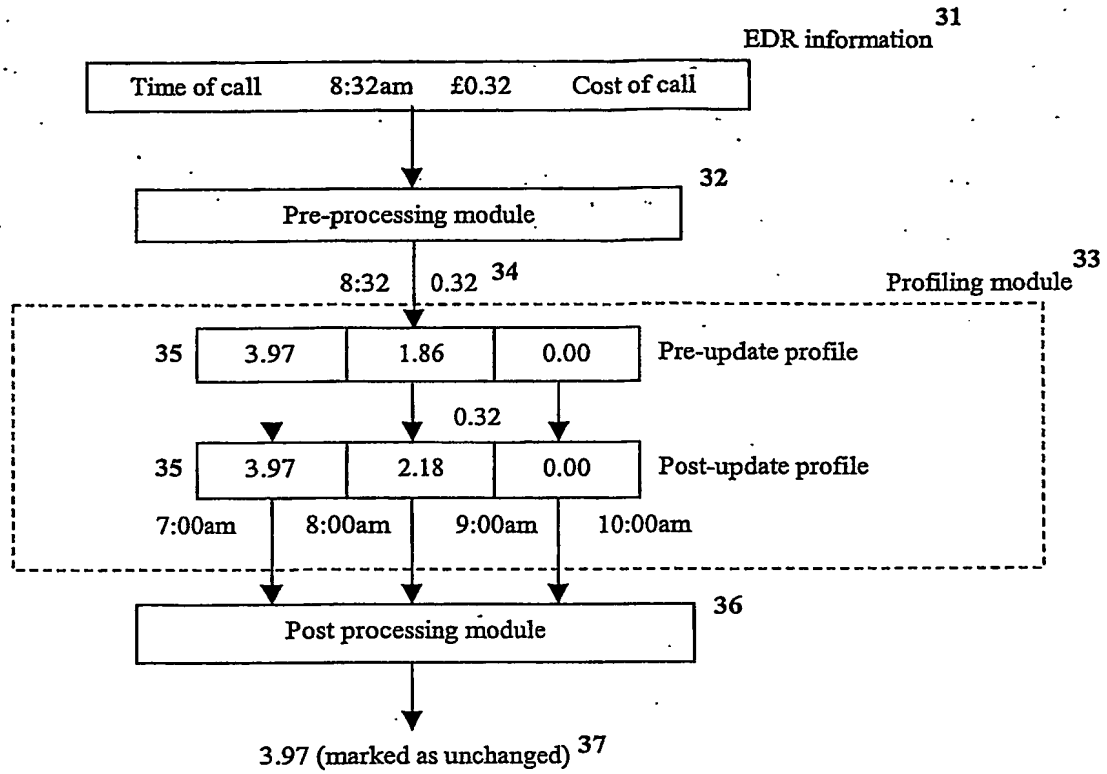


Fig. 3

BEST AVAILABLE COPY